JOURNAL OF COMBINATORIAL THEORY, Series B 40, 187–195 (1986)

An Exact Threshold Theorem for Random Graphs and the Node-Packing Problem

G. R. GRIMMETT

School of Mathematics, University of Bristol, Bristol, England

Communicated by the Editors

Received June 1, 1984

The usual linear relaxation of the node-packing problem contains no useful information when the underlying graph G has the property of "bicriticality." We consider a sparse random graph $G_n(m)$ obtained in the usual way from a random directed graph with fixed out-degree m and show that the probability that $G_n(2)$ is bicritical tends to $(1-2e^{-2})^{1/2}$ as $n \to \infty$. This confirms a conjecture by G. R. Grimmett and W. R. Pulleyblank (*Oper. Res. Lett.*, in press). © 1986 Academic Press, Inc.

1. THE RESULT

The node-packing problem on a graph G may be formulated as an integer linear program IP. In solving IP it is sometimes useful to have solved already the linear relaxation LR of IP, obtained by overlooking the conditions of integrality on the variables, and this approach was found by Nemhauser and Trotter [6]. Pulleyblank [7] has shown that the solution to LR contains no useful information about the solution to IP whenever G has the property of "bicriticality," and Grimmett and Pulleyblank [5] have shown that (in the context of a certain type of sparse random graph) almost all graphs with average node-degree of 6 or more are bicritical. In this note, we prove a conjecture of [5] dealing with the exact threshold for a random graph to be bicritical.

First we require some definitions. Let $G=(V, E)$ be an undirected graph. A subset A of nodes is called *stable* if no pair of nodes in A is adjacent in G; the *neighbour set* $N(A)$ of a subset A of nodes is given by

$$N(A)=\{b \in V \backslash A: b \text{ is adjacent to some } a \in A\}.$$

A pair (A, B) of subsets of V is called a *k-pair* if

$$A \cap B = \varnothing \qquad \text{and} \qquad |A| = |B| = k.$$

The k-pair (A, B) of G is called *bad* if A is stable and $N(A) \subseteq B$. We call the

0095-8956/86 $3.00

graph G *bicritical* if G contains no bad k-pair for any value $k = 1, 2, \ldots$; this definition of bicriticality differs from that used in [5] but is equivalent to it by Theorem 2.2 of [7].

We study the following type of random graph. Fix a positive integer m and construct a random directed graph $D_n(m)$ on the node set $\{1, 2, \ldots, n\}$ as follows: for each node $i \in \{1, 2, \ldots, n\}$, node i is joined to exactly m nodes of $\{1, 2, \ldots, n\} \setminus \{i\}$ by directed edges oriented away from i, these m nodes being chosen randomly with replacement and independently of each other. From the directed graph $D_n(m)$, we obtain an undirected graph $G_n(m)$ by deleting all orientations and allowing multiple edges to coalesce. Such random graphs have been the subject of much study recently (see the references in [5], and also [4]).

It is shown in [5] that, as $n \to \infty$,

$$\begin{aligned} P(G_n(m) \text{ is bicritical}) &\to 0 \qquad \text{if} \quad m = 1 \\ &\to 1 \qquad \text{if} \quad m \geqslant 3, \end{aligned}$$

and it is conjectured that

$$P(G_n(2) \text{ is bicritical}) \to (1 - 2e^{-2})^{1/2};$$

it is the purpose of this paper to prove this conjecture.

THEOREM. *As* $n \to \infty$, $P(G_n(2)$ *is bicritical*$) \to (1 - 2e^{-2})^{1/2}$.

This result is interesting not only because of its application to the related problem in operations research, but also because it is an exact threshold theorem for the random graph $G_n(m)$.

2. THE PROOF

If $G = (V, E)$ is a graph, then we call a bad k-pair (A, B) of G *reducible* if, for some l, there exists a bad l-pair (C, D) such that

$$1 \leqslant l < k, \qquad C \subseteq A, \qquad D \subseteq B, \tag{1}$$

and *irreducible* if no such bad l-pair exists.

Note that a graph G, with n nodes, is not bicritical if and only if G contains an irreducible bad k-pair for some k satisfying $1 \leqslant k \leqslant \frac{1}{2}n$. Let K be the (random) smallest value of k such that $G_n(2)$ contains an irreducible bad k-pair, with the convention that $K = \infty$ if $G_n(2)$ is bicritical (and hence contains no such pair). We shall prove the theorem by a series of lemmas.

LEMMA 1. *There exists α satisfying $0<\alpha<\frac{1}{2}$ such that*

$$P(n^{1/3}\leqslant K\leqslant \alpha n)\to 0 \qquad \text{as} \quad n\to\infty. \tag{2}$$

Proof. Fix k satisfying $n^{1/3}\leqslant k\leqslant \frac{1}{2}n$, where $n\geqslant 11$, and let $\pi=(A, B)$ be a k-pair and let B_k be the number of irreducible bad k-pairs in $G_n(2)$. Then

$$\begin{aligned} P(K=k) &\leqslant E(B_k) \\ &= \binom{n}{k}\binom{n-k}{k} P(\pi \text{ is irreducible and bad}) \\ &\leqslant \frac{n^{2k}}{k!^2} P(\pi \text{ is irreducible and bad}). \end{aligned} \tag{3}$$

Also

$$P(\pi \text{ is irreducible and bad}) = P(\pi \text{ is irreducible} \mid \pi \text{ is bad})\, P(\pi \text{ is bad}) \tag{4}$$

and it is easy to see that

$$\begin{aligned} P(\pi \text{ is bad}) &= \left(\frac{k}{n-1}\right)^{2k}\left(1-\frac{k}{n-1}\right)^{2(n-2k)} \\ &\leqslant \left(\frac{k}{n}\right)^{2k} \exp\left(\frac{2k}{n-1}-\frac{2k}{n}(n-2k)\right) \end{aligned} \tag{5}$$

since $1-x\leqslant e^{-x}$ for $x\geqslant 0$. Next we find an upper bound for $P(\pi$ is irreducible$\mid\pi$ is bad$)$. Let $U=\{u_1, u_2,\ldots, u_N\}$ be the nodes in B such that the edges of $D_n(2)$ emanating from U are not incident to nodes in A; we assume that $u_i<u_j$ if $i<j$. The size N of U is binomially distributed with parameters k and $(1-k/(n-1))^2$. But $k\leqslant\frac{1}{2}n$ and so N is no less (in distribution) than a binomially distributed random variable with parameters k and $\frac{1}{5}$, giving, by standard facts about the tail of the binomial distribution, that there exists a constant $\beta>0$, independent of k, such that, whatever the value of k,

$$P(N\leqslant \tfrac{1}{6}k)\leqslant e^{-\beta k}. \tag{6}$$

Suppose that $N>\frac{1}{6}k$. Consider the nodes $u_1, u_2,\ldots, u_r$, where $r=[\frac{1}{6}k]$ (where $[x]$ is the integer part of x) and let $d(u_i)$ be the number of edges in $D_n(2)$ which emanate from a node in A and are incident to u_i, for $1\leqslant i\leqslant r$. Conditional on the event $\{N>\frac{1}{6}k\}$ and the nodes $u_1, u_2,\ldots, u_r$ being given, it is the case that the random variables $d(u_1),\ldots, d(u_r)$ are distributed just as if $u_1,\ldots, u_r$ were the first r nodes (in lexicographic order, say) of B. Suppose now that we are given that $\pi=(A, B)$ is a bad k-pair; this assumption does not affect the above observation about the distribution of N. Then all of the

$2k$ edges in the original construction which emanate from nodes in A are incident to nodes in B. Hence

$$\begin{aligned}
&P(d(u_i) \geqslant 1 \text{ for all } 1 \leqslant i \leqslant r \mid E) \\
&\quad = P(d(u_1) \geqslant 1 \mid E)\, P(d(u_2) \geqslant 1 \mid E, d(u_1) \geqslant 1) \cdots \\
&\qquad \cdots P(d(u_r) \geqslant 1 \mid E, d(u_1) \geqslant 1, \ldots, d(u_{r-1}) \geqslant 1) \\
&\quad \leqslant \left(1 - \left(1 - \frac{1}{k}\right)^{2k}\right)^r \\
&\quad \leqslant (1 - e^{-3})^{k/7} \qquad (7)
\end{aligned}$$

for all large k, where $E = \{\pi \text{ is bad}, N > \frac{1}{6}k\}$. Equation (6) now gives that

$$\begin{aligned}
&P(\pi \text{ is irreducible} \mid \pi \text{ is bad}) \\
&\quad \leqslant P(\nexists b \in B \text{ such that } N(b) \cap A = \varnothing \mid \pi \text{ is bad}) \\
&\quad \leqslant P(d(u_i) \geqslant 1 \text{ for } 1 \leqslant i \leqslant r \mid E) + P(N \leqslant \tfrac{1}{6}k) \\
&\quad \leqslant (1 - e^{-3})^{k/7} + e^{-\beta k}
\end{aligned}$$

for all large n. Hence there exists $\gamma > 0$ such that, for all $n^{1/3} \leqslant k \leqslant \frac{1}{2}n$ and all large n,

$$P(\pi \text{ is irreducible} \mid \pi \text{ is bad}) \leqslant e^{-\gamma k}.$$

Substitute this with (5) into (4), and use (3) to obtain, for large n,

$$\begin{aligned}
P(K = k) &\leqslant \exp\left(-k\left(\gamma - \frac{2}{n-1} - \frac{4k}{n}\right)\right) \\
&\leqslant \exp\left(-k\left(\frac{1}{2}\gamma - \frac{4k}{n}\right)\right).
\end{aligned}$$

Let $\alpha = \gamma/9$ and suppose that $n^{1/3} \leqslant k \leqslant \alpha n$. Then

$$P(K = k) \leqslant \exp\left(-\frac{1}{18}\gamma k\right)$$

and

$$\begin{aligned}
P(n^{1/3} \leqslant K \leqslant \alpha n) &\leqslant \sum_{k = n^{1/3}}^{\alpha n} \exp\left(-\frac{1}{18}\gamma k\right) \\
&\to 0 \quad \text{as} \quad n \to \infty. \qquad \blacksquare
\end{aligned}$$

Lemma 2. *With α chosen as in Lemma* 1,

$$P(\alpha n \leqslant K \leqslant \tfrac{1}{2}n) \to 0 \qquad as \quad n \to \infty.$$

Proof. If (A, B) is a bad k-pair then A is stable and $|A| \geqslant |N(A)|$, and it follows that $(N(A), A)$ is a "matched pair" in the jargon of Frieze [4] before Lemma 2.7 of that paper. By Lemma 2.8 of [4], the probability that $G_n(2)$ contains a matched pair arising thus from a bad k-pair with $k \geqslant \alpha n$ tends to 0 as $n \to \infty$, and the result of the lemma follows. We note that Frieze's calculations relate to a slightly different random directed graph $D'_n(2)$ in which the neighbours of each node i are drawn with replacement from the whole vertex set $\{1, 2, \ldots, n\}$; thus $D'_n(2)$ may contain loops. There are at least two possible ways of dealing with this difference. The first is to repeat his calculations for our model. The second is to observe that, conditional on there being no loops, $D'_n(2)$ is stochastically the same as $D_n(2)$. Furthermore the number of loops in $D'_n(2)$ has asymptotically the Poisson distribution with parameter 2. Hence, writing P' for the probability measure associated with $D'_n(2)$, we have that

$$\begin{aligned} P'(\alpha n \leqslant K \leqslant \tfrac{1}{2}n) &\geqslant P'(\alpha n \leqslant K \leqslant \tfrac{1}{2}n \mid \text{no loops})\, P'(\text{no loops}) \\ &= P(\alpha n \leqslant K \leqslant \tfrac{1}{2}n)\, P'(\text{no loops}) \end{aligned}$$

giving that

$$\begin{aligned} P(\alpha n \leqslant K \leqslant \tfrac{1}{2}n) &\geqslant \frac{P'(\alpha n \leqslant K \leqslant \tfrac{1}{2}n)}{P'(\text{no loops})} \\ &\to 0 \qquad \text{as} \quad n \to \infty, \end{aligned}$$

since the numerator approaches 0 and the denominator approaches e^{-2}. ∎

It follows from Lemmas 1 and 2 that, with probability $1 - o(1)$, $G_n(2)$ is not bicritical if and only if $G_n(2)$ contains a bad k-pair for some $k < n^{1/3}$.

For the next lemma, we require another piece of notation. For any set A of nodes of $D_n(2)$, we write $D(A)$ for the subset of $\{1, 2, \ldots, n\} \backslash A$ containing endnodes of edges of $D_n(2)$ emanating from nodes in A.

Lemma 3. *Let E_k be the event that there exists a bad k-pair $\pi = (A, B)$ such that $D(B) \cap A \neq \varnothing$. Then*

$$P(E_k \textit{ occurs for some } k \textit{ satisfying } 1 \leqslant k \leqslant n^{1/3}) \to 0 \textit{ as } n \to \infty.$$

Proof. Straightforward calculations yield, for $k \leqslant n^{1/3}$,

$$P(E_k) \leqslant \binom{n}{k}\binom{n-k}{k}\left(\frac{k}{n-1}\right)^{2k}\left(1-\frac{k}{n-1}\right)^{2(n-2k)}\left(\frac{2k^2}{n-1}\right)$$
$$\leqslant \frac{\beta}{n}\frac{k^{2k+2}e^{-2k}}{(k!)^2}$$
$$\sim \frac{\gamma k}{n} \qquad \text{for large } k$$

where β and γ are constants. Hence, as $n \to \infty$,

$$P(E_k \text{ for some } k \text{ satisfying } 1 \leqslant k \leqslant n^{1/3}) \leqslant \frac{1}{n}\sum_{k=1}^{n^{1/3}} \delta k \to 0$$

where δ is a constant. ∎

Here is a final definition. Let G be a graph and let $d(v)$ be the degree of the node v of G. A *1-circuit* in G is a subgraph of G containing exactly two nodes u, v such that

(a) u and v are adjacent,

(b) either $d(u)=1$ or $d(v)=1$ or both.

A *k-circuit* in G is a subgraph G' of G such that

(c) G' is a circuit of length $2k$,

(d) if the nodes of G' are labelled in some cyclic order $c_1, c_2,\ldots, c_{2k}$ around this circuit, then either $d(c_1)=d(c_3)=\cdots=d(c_{2k-1})=2$ or $d(c_2)=d(c_4)=\cdots=d(c_{2k})=2$ or both.

Clearly $G_n(2)$ is not bicritical if it contains a k-circuit for some $k \geqslant 1$ (since, for example, if $c_1, c_2,\ldots, c_{2k}$ are the nodes of a k-circuit where $k>1$, taken in the cyclic ordering of (d) above with, say, $d(c_1)=d(c_3)=\cdots=d(c_{2k-1})=2$, then (A, B) is a bad k-pair where $A=\{c_1, c_3,\ldots, c_{2k-1}\}$ and $B=\{c_2, c_4,\ldots, c_{2k}\}$). It turns out that, with probability $1-o(1)$, $G_n(2)$ is not bicritical *if and only if* it contains a k-circuit for some k satisfying $1 \leqslant k \leqslant n^{1/3}$.

LEMMA 4. *Let C_n be the event that $G_n(2)$ contains a k-circuit for some k. Then*

$$P(\{G_n(2) \textit{ is not bicritical}\}\backslash C_n) \to 0 \quad \textit{as } n \to \infty.$$

Proof. By Lemmas 1, 2, and 3, with probability $1-o(1)$ it is the case that $G_n(2)$ is not bicritical if and only if it contains an irreducible bad k-pair $\pi=(A, B)$ such that $k \leqslant n^{1/3}$ and $D(B) \cap A = \varnothing$. Suppose that $\pi=(A, B)$ is such a k-pair with $k>1$, and let H be the bipartite graph with node sets A and B and with edge set the edges inherited from $G_n(2)$. Let $e_1, \ldots, e_{2k}$ be the edges of $D_n(2)$ which emanate from nodes in A. No pair (e_i, e_j) of edges is a pair of parallel edges, since if some node $a \in A$ is adjacent to only one node, b, say, of B then $(\{a\}, \{b\})$ is a bad 1-pair, contradicting the irreducibility of π. Hence H contains exactly $2k$ edges and the degree of each node $a \in A$ equals 2. Similarly the degree of each node $b \in B$ equals 2, since if this does not hold then there exists $b \in B$ with degree 1, adjacent in H to some node, a, say, of A implying that $(A \backslash \{a\}, B \backslash \{b\})$ is a bad $(k-1)$-pair and contradicting the irreducibility of π. Therefore all node-degrees in H equal 2, giving that H is the union of circuits which are both node-disjoint and edge-disjoint. There can only be one circuit in this union, since if there are two or more then π is reducible. Hence H is a circuit of length $2k$, and $G_n(2)$ contains a k-circuit. The case when $k=1$ is obvious. ∎

LEMMA 5. *There exists a constant η such that, for all $1 \leqslant L \leqslant n^{1/3}$ and all large n,*

$$P(G_n(2) \text{ contains a } k\text{-circuit with } L \leqslant k \leqslant n^{1/3}) \leqslant \eta L^{-1} + o(1)$$

where the o-term depends on n alone.

Proof. By the discussion before Lemma 4, each k-circuit C gives rise to a bad k-pair (A_C, B_C) (if C gives rise to two bad k-pairs, then we denote by (A_C, B_C) the outcome of a random choice of one of these two possibilities). If there exists a k-circuit C such that $1 \leqslant k \leqslant n^{1/3}$ and $D(B_C) \cap A_C \neq \varnothing$, then there exists a bad k-pair (A, B) with $1 \leqslant k \leqslant n^{1/3}$ and $D(B) \cap A \neq \varnothing$; by Lemma 3, this occurs with probability $o(1)$. Hence

$$P(G_n(2) \text{ contains a } k\text{-circuit with } L \leqslant k \leqslant n^{1/3}) \leqslant o(1) + \sum_{k=L}^{n^{1/3}} E(N_k)$$

where N_k is the number of k-circuits C in G with the property that $D(B_C) \cap A_C = \varnothing$. A simple counting argument shows that, for $k \leqslant n^{1/3}$,

$$\begin{aligned} E(N_k) &\sim \frac{n(n-1)\cdots(n-2k+1)}{2k}\left(\frac{2}{(n-1)^2}\right)^k\left(1-\frac{k}{n-1}\right)^{2(n-k)} \\ &\sim \frac{2^k e^{-2k}}{2k} \qquad \text{for large } n. \end{aligned} \tag{8}$$

Therefore

$$P(G_n(2) \text{ contains a } k\text{-circuit with } L \leqslant k \leqslant n^{1/3}) \leqslant \eta L^{-1} + o(1)$$

where $\eta = (1 - 2e^{-2})^{-1}$, for all large n. ∎

LEMMA 6. *Let M_k be the number of k-circuits in $G_n(2)$ for $1 \leqslant k \leqslant \frac{1}{2}n$. For all fixed $L \geqslant 1$, the family $(M_1, M_2, \dots, M_L)$ is asymptotically jointly distributed as independent random variables, M_k having, asymptotically, a Poisson distribution with parameter*

$$\lambda_k = \frac{1}{2k} 2^k e^{-2k}$$

for $k = 1, 2, \dots, L$. That is to say, for all $(m_1, m_2, \dots, m_L) \in \mathbb{N}^L$,

$$P(M_1 = m_1, \dots, M_L = m_L) \to \prod_{k=1}^{L} \frac{1}{m_k!} \lambda_k^{m_k} e^{-\lambda_k} \quad \text{as } n \to \infty.$$

Proof. Fix k satisfying $1 \leqslant k \leqslant L$. By the argument in the proof of Lemma 5, there is probability $o(1)$ that there exists a k-circuit C for which $D(B_C) \cap A_C \neq \varnothing$. Hence we may restrict ourselves to studying the number N_k of k-circuits C such that $D(B_C) \cap A_C = \varnothing$. By (8)

$$E(N_k) \to \lambda_k \qquad \text{as} \quad n \to \infty.$$

The Poisson limit theorem of Lemma 6 now follows in exactly the standard fashion, and we omit the details of the proof, which is tedious but routine (see Erdös and Rényi [3, p. 27], Schürger [8, p. 50], and Bollobás [1, p. 201; 2, p. 92]). ∎

At last we are ready to complete the proof of the theorem. From Lemmas 4, 5, and 6,

$$\begin{aligned}
&P(G_n(2) \text{ is not bicritical}) \\
&\quad = o(1) + P(G_n(2) \text{ contains a } k\text{-circuit for some } k) \\
&\quad = o(1) + O(L^{-1}) + P(G_n(2) \text{ contains a } k\text{-circuit for some } 1 \leqslant k \leqslant L) \\
&\quad = o(1) + O(L^{-1}) + 1 - P(M_1 = M_2 = \cdots = M_L = 0) \\
&\quad = o(1) + O(L^{-1}) + 1 - \exp\left(-\sum_{k=1}^{L} \lambda_k\right) \qquad (9)
\end{aligned}$$

where the o-terms and the O-terms depend, respectively, on n and L only. Let $n \to \infty$ and then $L \to \infty$ in (9), in that order, to obtain

$$P(G_n(2) \text{ is not bicritical}) \to 1 - \exp\left(-\sum_{k=1}^{\infty} \frac{1}{2k}(2e^{-2})^k\right)$$
$$= 1 - (1 - 2e^{-2})^{1/2}$$

as required.

Acknowledgment

I am grateful to Alan Frieze for discussing the proof of Lemma 2 with me.

References

1. B. Bollobás, Threshold functions for small subgraphs, *Math. Proc. Cambridge Philos. Soc.* **90** (1981), 197–206.
2. B. Bollobás, Random graphs, *in* "Combinatorics" (H. N. V. Temperley, Ed.), London Mathematical Society Lecture Notes 52, pp. 80–102, 1981.
3. P. Erdös and A. Rényi, On the evolution of random graphs, *Publ. Math. Inst. Hungar. Acad. Sci.* **5A** (1960), 17–61.
4. A. M. Frieze, Maximum matchings in a class of random graphs, *J. Combin. Theory Ser. B* **40** (1986), 196–212.
5. G. R. Grimmett and W. R. Pulleyblank, Random near-regular graphs and the node packing problem, *Oper. Res. Lett.* in press.
6. G. L. Nemhauser and L. E. Trotter, Vertex packings: Structural properties and algorithms, *Math. Programming* **8** (1975), 232–248.
7. W. R. Pulleyblank, Minimum node covers and 2-bicritical graphs, *Math. Programming* **17** (1979), 91–103.
8. K. Schürger, Limit theorems for complete subgraphs of random graphs, *Period. Math. Hungar.* **10** (1979), 47–53.